

5' -> 3' direction of transcription

5'	9	18	27	36	45	54
	GG GGA AGC AGC	TGG AGT GCG ACC	GCC ACC GCA GCC	ACC CTG CAA CCG	CCA GTC	
	63	72	81	90	99	108
	GGA GGT GCA GTC	CGT AGG CCC TGG CCC	CCG GGT GGG CCC	TTG GGG AGT CGG	CGC	
	117	126	135	144	153	162
	CGC TCC CGA GGA	GCT GCA AGG CTC	GCC CCT GCC CGG	CGT GGA GGG CGC	GGG	
	171	180	189	198	207	216
	CGC GGA GAA AGT	GAA AGG AAA TTG	GAA AAT TGT GAG	TGG ACC TTC	TGA TAC	
	225	234	243	252	261	270
	TGC TCC TCC TTG	CGT GGA AAA GGG	GAA AGA ACT GCA	TGC ATA TTA	TTC AGC	GTC
	279	288	297	306	315	324
	CTA TAT TCA AAG	GAT ATT CTT GGT	GAT CTT GGA AGT	GTC CGT ATC	ATG GAA	TCA

						M E S
	333	342	351	360	369	378
	ATC TCT ATG ATG	GGA AGC CCT AAG	AGC CTT AGT GAA	ACT TGT TTA	CCT AAT	GGC
	---	---	---	---	---	---
I	S	M	G	S	P	K
						S
						L
						C
						T
						E
						S
						N
						G

FIGURE 1A

387	ATA	GGT	ATC	AAA	GAT	GCA	AGG	AAG	GTC	ACT	GTA	GGT	GTG	ATT	GGA	AGT	GGA
	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
	I	N	G	I	K	D	A	R	K	V	T	V	G	V	I	G	S
																	G
441	GAT	TTT	GCC	AAA	TCC	TTG	ACC	ATT	CGA	CTT	ATT	AGA	TGC	GGC	TAT	CAT	GTG
	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
	D	F	A	K	S	L	T	I	R	L	I	R	C	G	Y	H	V
																	V
495	ATA	GGA	AGT	AGA	AAT	CCT	AAG	TTT	GCT	TCT	GAA	TTT	TTT	CCT	CAT	GTG	GTA
	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
	I	G	S	R	N	P	K	F	A	S	E	F	F	P	H	V	D
																	V
549	GTC	ACT	CAT	CAT	GAA	GAT	GCT	CTC	ACA	AAA	ACA	AAT	ATA	ATA	TTT	GTT	GCT
	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
	V	T	H	H	E	D	A	L	T	K	T	N	I	I	F	V	A
																	I
603	CAC	AGA	GAA	CAT	TAT	ACC	TCC	CTG	TGG	GAC	CTG	AGA	CAT	CTG	CTT	GTG	GGT
	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
	H	R	E	H	Y	T	S	L	W	D	L	R	H	L	L	V	G
																	K

FIGURE 1B

657	ATC	CTG	ATT	GAT	GTG	AGC	AAT	AAC	ATG	AGG	ATA	AAC	CAG	TAC	CCA	GAA	TCC	AAT	702
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
I	L	I	D	V	S	N	N	N	M	R	I	N	Q	Y	P	E	S	N	
711	GCT	GAA	TAT	TTG	GCT	TCA	TTA	TTC	CCA	GAT	TCT	TTG	ATT	GTC	AAA	GGA	TTT	AAT	756
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
A	E	Y	L	A	S	L	F	P	D	S	L	I	V	K	G	F	N		
765	GTT	GTC	TCA	GCT	TGG	GCA	CTT	CAG	TTA	GGA	CCT	AAG	GAT	GCC	AGC	CGG	CAG	GTT	810
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
V	S	A	W	A	L	Q	L	G	P	K	D	A	S	R	Q	V			
819	TAT	ATA	TGC	AGC	AAC	AAT	ATT	CAA	GCG	CGA	CAA	CAG	GTT	ATT	GAA	CTT	GCC	CGC	864
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Y	I	C	S	N	N	I	Q	A	R	Q	Q	V	I	E	L	A	R		
873	CAG	TTG	AAT	TTC	ATT	CCC	ATT	GAC	TTG	GGA	TCC	TTA	TCA	TCA	GCC	AGA	GAG	ATT	918
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Q	L	N	F	I	P	I	D	L	G	S	L	S	S	A	R	E	I		

FIGURE 1C

927	936	945	954	963	972
GAA AAT TTA CCC CTA CGA CTC TTT ACT CTC TGG AGA GGG CCA GTG GTG GTA GCT					
---	---	---	---	---	---
E N L P L R L F T L W R G P V A					
981	990	999	1008	1017	1026
ATA AGC TTG GCC ACA TTT TTT TTC CTT TAT TCC TTT GTC AGA GAT GTG ATT CAT					
---	---	---	---	---	---
I S L A T F F F L Y S F V R D V I H					
1035	1044	1053	1062	1071	1080
CCA TAT GCT AGA AAC CAA CAG AGT GAC TTT TAC AAA ATT CCT ATA GAG ATT GTG					
---	---	---	---	---	---
P Y A R N Q Q Q S D F Y K I P I E I V					
1089	1098	1107	1116	1125	1134
AAT AAA ACC TTA CCT ATA GTT GCC ATT ACT TTG CTC TCC CTA GTA TAC CTC GCA					
---	---	---	---	---	---
N K T L P I V A I T L L S L V Y L A					
1143	1152	1161	1170	1179	1188
GGT CTT CTG GCA GCT GCT TAT CAA CTT TAT TAC GGC ACC AAG TAT AGG AGA TTT					
---	---	---	---	---	---
G L L A A A Y Q L Y Y G T K Y R R F					

FIGURE 1D

1197	1206	1215	1224	1233	1242
CCA CCT TGG TTG GAA ACC TGG TTA CAG TGT AGA AAA CAG CTT GGA TTA CTA AGT					
---	---	---	---	---	---
P P W L E T W L Q C R K Q L G L L S					
1251	1260	1269	1278	1287	1296
TTT TTC TTC GCT ATG GTC CAT GTT GCC TAC AGC CTC TGC TTA CCG ATG AGA AGG					
---	---	---	---	---	---
F F F A M V H V A Y S L C L P M R R					
1305	1314	1323	1332	1341	1350
TCA GAG AGA TAT TTG TTT CTC AAC ATG GCT TAT CAG CAG GTT CAT GCA AAT ATT					
---	---	---	---	---	---
S E R Y L F L N M A Y Q Q V H A N I					
1359	1368	1377	1386	1395	1404
GAA AAC TCT TGG AAT GAG GAA GAA GTT TGG AGA ATT GAA ATG TAT ATC TCC TTT					
---	---	---	---	---	---
E N S W N E E E V W R I E M Y I S F					
1413	1422	1431	1440	1449	1458
GGC ATA ATG AGC CTT GGC TTA CTT TCC CTC CTC CTG GCA GTC ACT TCT ATC CCT TCA					
---	---	---	---	---	---
G I M S L G L L S L L A L A V T S I P S					

FIGURE 1E

1467	1476	1485	1494	1503	1512
GTG AGC AAT GCT TTA AAC TGG AGA GAA TTC AGT TTT ATT CAG TCT ACA CTT GGA					
V S N A L N W R E F S F I Q S T L G					
1521	1530	1539	1548	1557	1566
TAT GTC GCT CTG CTC ATA AGT ACT TTC CAT GTT TTA ATT TAT GGA TGG AAA CGA					
Y V A L L I S T F H V L I Y G W K R					
1575	1584	1593	1602	1611	1620
GCT TTT GAG GAA GAG TAC TAC AGA TTT TAT ACA CCA CCA AAC TTT GTT CTT GCT					
A F E E E Y Y R F Y T P P N F V L A					
1629	1638	1647	1656	1665	1674
CTT GTT TTG CCC TCA ATT GTA ATT CTG GGT AAG ATT ATT TTA TTC CTT CCA TGT					
L V L P S I V I L G K I I L F L P C					
1683	1692	1701	1710	1719	1728
ATA AGC CGA AAG CTA AAA CGA ATT AAA AAA GGC TGG GAA AAG AGC CAA TTT CTG					
I S R K L K R I K K K G W E K S Q F L					

FIGURE 1F

1737 1746 1755 1764 1773 1782
 GAA GAA GGT ATT GGA GGA ACA ATT CCT CAT GTC TCC CCG GAG AGG GTC ACA GTA
 --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
 E E G I G G T I P H V S P E R V T V
 1791 1800 1809 1818 1827 1836
 ATG TGA TGA TAA ATG GTG TTC ACA GCT GCC ATA TAA AGT TCT ACT CAT GCC ATT
 --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
 M
 1845 1854 1863 1872 1881 1890
 ATT TTT ATG ACT TCT ACG TTC AGT TAC AAG TAT GCT GTC AAA TTA TCG TGG GTT
 GA 3'
 --

FIGURE 1G

1	MES	I	S	M	G	S	P	K	S	L	S	E	T	C	L	P	N	G	I	N	G	I	K	D	A	R	K	7492448			
1	MES	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	R	K	g6572948			
31	V	T	V	G	V	I	G	S	G	D	F	A	K	S	L	T	I	R	L	I	R	C	G	Y	H	V	I	G	S	7492448	
6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	g6572948		
61	R	N	P	K	F	A	S	E	F	F	P	H	V	D	V	T	H	H	E	D	A	L	T	K	T	N	I	I	F	7492448	
6	-	-	-	-	-	-	-	-	-	-	-	-	-	D	I	T	N	Q	E	E	-	-	-	-	-	-	-	-	g6572948		
91	V	A	I	H	R	E	H	Y	T	S	L	W	D	L	R	H	L	L	V	G	K	I	L	I	D	V	S	N	M	7492448	
13	-	-	-	-	-	-	-	-	-	-	L	W	K	M	K	P	-	-	-	-	-	-	-	-	-	R	R	N	L	g6572948	
121	R	I	N	Q	Y	P	E	S	N	A	E	Y	L	A	S	L	F	P	D	S	L	I	V	K	G	F	N	V	S	7492448	
23	E	E	D	D	Y	L	H	K	D	T	G	-	E	T	S	M	L	K	R	P	V	L	-	-	-	-	-	-	-	g6572948	
151	A	W	A	L	Q	L	G	P	K	D	A	S	R	Q	V	Y	I	C	S	N	N	I	Q	A	R	Q	Q	V	I	E	7492448
44	-	-	-	L	H	L	-	H	Q	T	A	H	A	D	E	F	D	C	P	S	E	L	Q	H	T	Q	E	-	-	-	g6572948
181	L	A	R	Q	L	N	F	I	P	I	D	L	G	S	L	S	S	A	R	E	I	E	N	L	P	L	R	L	F	T	7492448
67	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	L	F	P	g6572948	
211	L	W	R	G	P	V	V	A	I	S	L	A	T	F	F	F	L	Y	S	F	V	R	D	V	I	H	P	Y	A	7492448	
70	Q	W	H	L	P	I	K	I	A	A	I	A	S	L	T	F	L	Y	T	L	L	R	E	V	I	H	P	L	A	g6572948	

FIGURE 2A

241	R	N	Q	Q	S	D	F	Y	K	I	P	I	E	I	V	N	K	T	L	P	I	V	A	I	T	L	L	S	L	V	7492448
100	T	S	H	Q	Q	Y	F	Y	K	I	P	I	L	V	I	N	K	V	L	P	M	V	S	I	T	L	L	A	L	V	g6572948
271	Y	L	A	G	L	L	A	A	A	Y	Q	L	Y	Y	G	T	K	Y	R	R	F	P	P	W	L	E	T	W	L	Q	7492448
130	Y	L	P	G	V	I	A	A	I	V	Q	L	H	N	G	T	K	Y	K	K	F	P	H	W	L	D	K	W	M	L	g6572948
301	C	R	K	Q	L	G	L	L	S	F	F	F	A	M	V	H	V	A	Y	S	L	C	L	P	M	R	R	S	E	R	7492448
160	T	R	K	Q	F	G	L	L	S	F	F	F	A	V	L	H	A	I	Y	S	L	S	Y	P	M	R	R	S	Y	R	g6572948
331	Y	L	F	L	N	M	A	Y	Q	Q	V	H	A	N	I	E	N	S	W	N	E	E	E	V	W	R	I	E	M	Y	7492448
190	Y	K	L	L	N	W	A	Y	Q	Q	V	Q	Q	N	K	E	D	A	W	I	E	H	D	V	W	R	M	E	I	Y	g6572948
361	I	S	F	G	I	M	S	L	G	L	L	S	L	L	A	V	T	S	I	P	S	V	S	N	A	L	N	W	R	E	7492448
220	V	S	L	G	I	V	G	L	A	I	L	A	L	L	A	V	T	S	I	P	S	V	S	D	S	L	T	W	R	E	g6572948
391	F	S	F	I	Q	S	T	L	G	Y	V	A	L	L	I	S	T	F	H	V	L	I	Y	G	W	K	R	A	F	E	7492448
250	F	H	Y	I	Q	S	K	L	G	I	V	S	L	L	L	G	T	I	H	A	L	I	F	A	W	N	K	W	I	D	g6572948
421	E	E	Y	Y	R	F	Y	T	P	P	N	F	V	L	A	L	V	L	P	S	I	V	I	L	G	K	I	L	L	F	7492448
280	I	K	Q	F	V	W	Y	T	P	P	T	F	M	I	A	V	F	L	P	I	V	V	L	I	F	K	S	I	L	F	g6572948

FIGURE 2B

451	L	P	C	I	S	R	K	L	K	R	I	K	K	G	W	E	K	S	Q	F	L	E	F	G	I	G	G	T	I	P	7492448		
310	L	P	C	L	R	K	K	I	L	K	I	R	H	G	W	E	D	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	g6572948
481	H	V	S	P	E	R	V	-	T	V	M																	7492448					
329	K	I	N	K	T	E	I	C	S	Q	L																	g6572948					

FIGURE 2C